Feasibility of a 100 year Reanalysis Using Only Surface Pressure Data

Gilbert P. Compo, Jeffrey S. Whitaker,

and Prashant D. Sardeshmukh

NOAA-CIRES Climate Diagnostics Center

University of Colorado at Boulder

Accepted in the Bulletin of the American Meteorological Society

23 September 2005

Corresponding author: Dr. Gilbert P. Compo, NOAA-CIRES Climate Diagnostics Center R/CDC1, 325 Broadway, Boulder, CO 80305-3328. (303) 497-6115. compo@colorado.edu

Abstract

Climate variability and global change studies are increasingly focused on understanding and predicting regional changes of daily weather statistics. Assessing the evidence for such variations over the last hundred years requires a daily tropospheric circulation dataset. The only dataset available for the early 20th century consists of error-ridden hand-drawn analyses of the mean sea level pressure field over the Northern Hemisphere. Modern data assimilation systems have the potential to improve upon these maps, but prior to 1948, few digitized upper-air sounding observations are available for such a "reanalysis." We investigate the possibility that the additional number of newly recovered surface pressure observations is sufficient to generate useful weather maps of the lower-tropospheric extratropical circulation back to 1890 over the Northern Hemisphere, and back to 1930 over the Southern Hemisphere. Surprisingly, we find that using an advanced data assimilation system based on an ensemble Kalman filter, it would be feasible to produce high-quality maps of even the upper troposphere using only surface pressure observations. For the beginning of the 20th century, the errors of such upper-air circulation maps over the Northern Hemisphere in winter would be comparable to the two to three day errors of modern weather forecasts.

Capsule

A modern data assimilation system could use the available surface pressure observations to produce a consistent dataset of the daily extratropical circulation from the late 19th century to the present.

In the early 1940s, officials in the United States recognized the need for a comprehensive series of synoptic weather maps extending over the 20th century to use for meteorological studies and further the war effort (United States Weather Bureau 1944). The Weather Bureau and U.S. Army keyed in manuscript station and marine observations and trained meteorologists to make daily sea level pressure (SLP) analyses, resulting in the US Historical Weather Map Series (USHWM) spanning January 1899 to June 1939. An example map is shown in Fig. 1. These hand-drawn analyses were recognized for their value, and were digitized in the 1970s by the National Weather Service Extended Forecast Division (NCAR 2005). This original "reanalysis" effort and subsequent additions form the Daily Northern Hemisphere Sea Level Pressure Grids dataset.[1] Not long after their digitization, however, researchers realized that the dataset contains substantial errors and inhomogeneities (e.g. Williams and van Loon 1976, Trenberth and Paolino 1980, Hayden 1984, Jones 1987). Further, the maps do not make use of all observations, being limited to those collections that were archived by the United States prior to the 1940s and then to those observations collected in near-real time. These errors and limitations hamper the applicability of this dataset to studies of climate variability and change. The surface nature of the maps further limits their use. Climate studies need four dimensional, dynamically consistent circulation fields to make progress.

In the 1980's, the operational analyses produced by national meteorological centers to support numerical weather prediction (NWP) were recognized as being difficult to use for climate studies because they were based on observations, data assimilation algorithms, and models that were changing throughout the record as the centers worked to improve their forecast skill (Bengtsson and Shukla 1988). It was proposed that a "reanalysis" be performed, whereby a fixed NWP model and data assimilation algorithm would be used to blend the historical

[1]These grids are archived at the National Center for Atmospheric Research (NCAR) as ds010.0 and are available online at http://dss.ucar.edu/datasets/ds010.0/.

observations into a consistent set of analyses needed for climate research. This reanalysis methodology was expected to eliminate the spurious discontinuities present in the operational analyses. Reanalyses such as the National Centers for Environmental Prediction – National Center for Atmospheric Research (NCEP-NCAR) reanalysis (Kalnay et. al. 1996, Kistler et. al. 2001) and the European Center for Medium Range Weather Forecasting (ECMWF) 40 year reanalysis (Simmons and Gibson 2000) have now been carried out by several centers. These datasets have been used in many studies of atmospheric variability, particularly on synoptic to interannual timescales.

In addition to reanalyses, several efforts have been successful at reconstructing global fields of monthly mean surface temperature (Kaplan et. al. 1998, Rayner et. al. 2003, Smith and Reynolds 2004a) and sea level pressure (Smith and Reynolds 2004b, Kaplan et. al. 2000) over the past 150 years using statistical methods. These longer datasets have been used in many studies of seasonal to decadal variability and climate change.

Three issues limit one's ability to examine climate variability and climate change from synoptic to decadal timescales over the last century. The reanalysis datasets are only available beginning in the latter half of the 20th century. The reconstructed datasets are only at the surface and only monthly means. Monthly means, while valuable in many respects, are unable to capture either high-frequency, synoptic-scale weather events (such as intense extratropical cyclones), or lower-frequency, planetary-scale climate events (such as blocking or the Madden-Julian Oscillation) that are of interest in climate research. The only available daily dataset for the first half of the 20th century, the Northern Hemisphere SLP Grids, is also only of a surface field and has well-documented problems.

One might consider reanalyzing the available historical data to extend the record. Unfortunately, the modern reanalysis methodology of using all available observations has resulted in significant inhomogeneities from the time-changing observation network. These

range from understated storm track variability (Harnik and Chang 2003, Chang and Fu 2003, Hodges et. al. 2003) to incorrect tropical variability (Newman et. al. 2000) to spurious long-term trends (Trenberth and Smith 2005, Bengtsson et. al. 2004a, Bengtsson et. al. 2004b, Kinter et. al. 2004, Kistler et. al. 2001, Basist and Chelliah 1997). As one attempts to reanalyze the early 20th century, or even the 19th century, scant upper-air data will be available (Bronnimann et. al. 2005). Recent results suggest that the current generation of reanalyses depends on the upper-air data (Bengtsson et. al. 2004b, Kanamitsu and Hwang, 2005) to produce reasonable tropospheric fields.

There are new research efforts to remedy this situation. Aided by the large quantity of newly recovered pressure observations, the European Multidecadal EMULATE project has used statistical reconstruction to create 150 years of daily sea level pressure maps over the data-rich North Atlantic and Europe (Ansell et. al. 2005, submitted to J. Climate). The question remains open whether modern data assimilation systems can derive useful information, not just of the near-surface daily atmospheric circulation, but also in the free troposphere from only surface observations. Results to date are mixed.

Whitaker et. al. (2003, 2004) used real surface pressure observations in a network simulating the 1915 network and an ensemble Kalman filter to show that a useful analysis of the 4-times daily atmospheric circulation at the surface and at 500 hPa could be created for the Northern Hemisphere. Anderson et. al. (2005) using a different ensemble filter in a perfect model experiment, and Simmons and Thepaut (2003, personal communication) using the ECMWF four-dimensional variational assimilation (4DVAR) system with a modern network of real observations, achieved results similar to Whitaker et. al. In contrast, using three-dimensional variational assimilation (3DVAR) systems, Bengtsson et. al. (2004b) and Kanamitsu and Hwang (2005) have reported that surface pressure observations alone cannot be used to make a good analysis of mid-tropospheric circulations, even in the relatively data-rich Northern

Hemisphere. In both of these 3DVAR studies, however, the relative weighting of the first-guess forecast and the observations was that appropriate for the modern, complete observing system. These weights may not be appropriate when only sparse observations are assimilated.

Recently, national efforts such as the Climate Change Science Program (Mahoney et. al. 2003) and the Ongoing Analysis of the Climate System (Arkin et. al. 2003) and international efforts such as the Global Climate Observing System (Mason et. al. 2004) and Global Earth Observation System of Systems (GEO 2005) have called for reanalysis datasets extending as far back as possible to compare the patterns and magnitudes of recent and projected climate changes in means and extremes with past changes. It is hoped that these longer reanalysis datasets enable researchers to address issues such as the range of natural variability of extreme events (e.g., floods, droughts, hurricanes, extratropical cyclones, and cold waves) and how El Nino/Southern Oscillation and other climate modes alter these events. Given the conflicting previous results, it has not been firmly established how useful such reanalyses based only on surface observations would be to address these issues.

In this study, we systematically examine the potential for data assimilation systems to reanalyze not only the near surface circulation, but also the upper-air circulation from the late 19th century to present using only surface pressure observations. We have chosen three assimilation schemes to make our assessment. A 3DVAR scheme very similar to that used for the NCEP-NCAR reanalysis will allow us to test a system that has been extensively used and studied. An ensemble Kalman filter will represent the potential for advanced data assimilation systems to improve upon older 3DVAR systems. As a baseline measure, similar to using climatology for assessing skill in probabilistic forecasting, we will use a climatologically based statistical interpolation scheme with no dynamical model to advance information to the next analysis time step. This will enable us to quantify the importance of propagating information with a dynamical model.

In all of our experiments we assimilate only surface pressure observations. Measurements of surface pressure have been made consistently throughout the period and should provide more information about the large-scale free tropospheric circulation than surface wind or surface temperature. The surface pressure information, through geostrophy, can be expected to yield a reasonable approximation to the barotropic part of the flow, which accounts for a substantial part of the total flow. Idealized studies have shown that surface pressure footprints of vertically tilted baroclinic modes contain useful information about the amplitudes and phases of those modes throughout the troposphere (Bengtsson 1980).

Simulating past surface pressure networks.

To investigate the feasibility of a 100 year or longer reanalysis, we reduced the 2001 observational network to resemble that available historically (see Sidebar 1) and then performed a series of observing system experiments, assimilating these reduced networks with different data assimilation systems.

We chose to simulate the observational densities of four representative 5-year periods centered on 1895, 1905, 1915, and 1935 to assess the expected analysis quality from the late 19th century to the early 20th century. To create spatial and temporal distributions using 2001 data that approximate the data density in these periods (Fig. SB1), we first eliminated all observations from the quality-controlled observational database used to generate the NCEP-NCAR reanalyses (Kalnay et. al. 1996) except the surface pressure observations from radiosonde reports and surface marine reports within 30 minutes of the synoptic time. These criteria reduced the total number of observations from 150,000 per analysis to a 2001 thinned network of 2,000 surface pressure observations per analysis. We randomly reduced the 2001 thinned network further so that the number of observations retained in 5°X5° regions over the globe mimicked the number of observations actually available per day in each of our historical periods. Ideally, we would specify the observational error statistics appropriate for our historical periods. However, at

present, nothing is known about these statistics. Therefore, we decided to use the modern observational estimates of 1 hPa for station observations and 1.6 hPa for marine observations. An example analysis network for the 1895 simulation is shown in Fig. 2. In the example, the sparse network density and the paucity of observations available over the central Pacific are especially apparent.

With the four sets of surface pressure observations, we investigated the ability of our three candidate assimilation systems to create 6-hourly three-dimensional analyses for June and December 2001: the NCEP Climate Data Assimilation System (CDAS, Kistler et. al., 2001; Kalnay et. al., 1996) modified for surface pressure only (CDAS-SFC), an Ensemble Kalman Filter or Ensemble Filter (EnsFilt, Whitaker et. al. 2004), and a "Climatological Ensemble Filter" (EnsClim, Whitaker et. al. 2004) loosely approximating a traditional statistical interpolation scheme. The NCEP CDAS system is a 3DVAR system (Parrish and Derber 1992) with static uncertainties (background-error covariances) for its first guess forecast field tuned to represent the uncertainties of 6-hour model forecasts started from analyses that used all of the observations in the modern network (Kistler et. al. 2001). We modified the CDAS system in several ways to adapt it to the assimilation of only surface pressure observations, forming the CDAS-SFC system, as described in Whitaker et. al. (2004). We inflated the background error covariances to reflect the less accurate analyses and subsequent forecasts expected when only assimilating surface pressure observations. Through trial and error we determined that multiplying by a factor of 16 was necessary to reduce the analysis error, but we did not modify the spatial structure of the covariances. This changed the expected error in the first guess from about 2 hPa in mid-latitudes to about $2 \times \sqrt{16} = 8\,\text{hPa}$. We also found that a constraint on the divergence tendency in the CDAS system severely limited the effect of a surface pressure observation on the analysis. This constraint was turned off with no detrimental effects and resulted in a significant decrease in analysis error.

Idealized studies have suggested that an Ensemble Kalman Filter (Evensen 1994, Burgers et. al. 1998), or Ensemble Filter, may have smaller analysis errors than 3DVAR when observations are sparse (Hamill and Snyder 2000). To evaluate this potential advantage, we have used a particular Ensemble Filter (Whitaker and Hamill 2002) from the family of "Ensemble Square Root Filters" recently reviewed by Tippett et. al. (2003).

This Ensemble Filter has a simple implementation when observations are processed one at a time. Assume that we have an *n*-member ensemble of first guess fields, with the *j*th member $\mathbf{x}_j^{\mathbf{b}}$ representing the complete state vector of the forecast model (e.g., wind, temperature, humidity, and surface pressure fields on the model domain). The sample mean of these fields is $\overline{\mathbf{x}}^{\mathbf{b}} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{x}_j^{\mathbf{b}}$, and the deviations from the mean are $\mathbf{x}'^{\mathbf{b}}_j = \mathbf{x}_j^{\mathbf{b}} - \overline{\mathbf{x}}^{\mathbf{b}}$. We denote the first surface pressure observation to be assimilated as y^o and the ensemble mean and deviations interpolated to the observation location as $\overline{y_j}^{b} = \mathbf{H}\overline{\mathbf{x}_j}^{\mathbf{b}}$ and $y'^{b}_j = \mathbf{H}\mathbf{x}'^{\mathbf{b}}_j$, respectively, with $\mathbf{H}$ being the operator that interpolates the first guess surface pressure field to the observation location. We combine the first guess ensemble and the observation to form an *n*-member analysis ensemble, whose mean $\overline{\mathbf{x}}^{\mathbf{a}}$ and deviations $\mathbf{x}'^{\mathbf{a}}_j$ are calculated via

$$\overline{\mathbf{x}}^{\mathbf{a}} = \overline{\mathbf{x}}^{\mathbf{b}} + \mathbf{K}\left(y^o - \overline{y}^{b}\right) \tag{1}$$

and

$$\mathbf{x}'^{\mathbf{a}}_j = \mathbf{x}'^{\mathbf{b}}_j - \tilde{\mathbf{K}}\left(y'^{b}_j\right) \tag{2}$$

where the Kalman gain $\mathbf{K}$ is given by

$$\mathbf{K} = \frac{1}{n-1}\sum_{j=1}^{n} \mathbf{x}'^{\mathbf{b}}_j y'^{b}_j \left(\frac{1}{n-1}\sum_{j=1}^{n} y'^{b}_j y'^{b}_j + R\right)^{-1}, \tag{3}$$

and the modified Kalman gain $\tilde{\mathbf{K}}$ is given by

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{R}{\frac{1}{n-1}\sum_{j=1}^{n} y'^{b}_{j} y'^{b}_{j} + R}} \right)^{-1} \mathbf{K} \quad . \tag{4}$$

We assume that the observation y^o has a specified observational error variance R that represents both the measurement error associated with the observation and the error associated with representing a large area (i.e., an NWP model grid box) from a point measurement. We further assume that the error in y^o is uncorrelated with all the other observations to be assimilated. The model means, variances, and co-variances in eqs 1-4 are all unbiased sample estimates from the n=100 member ensemble. Within the limitations of using an imperfect model and finite ensembles, this formulation represents a minimum-error estimate of the "true" state (Lorenc 1986), represented here by the analysis ensemble mean $\overline{\mathbf{x}}^{\mathbf{a}}$. To assimilate subsequent observations, the 100 members of the analysis ensemble $\mathbf{x}^{\mathbf{a}}_j$ become the new first guess ensemble and equations 1-4 are applied iteratively for each observation. After all available observations have been assimilated, the 100 member set of analyses $\mathbf{x}^{\mathbf{a}}_j = \overline{\mathbf{x}}^{\mathbf{a}} + \mathbf{x}'^{\mathbf{a}}_j$ becomes the 100 initial conditions for the subsequent 6-hour forecast/analysis cycle (see Whitaker et. al. 2004 for a complete description).

Model error and sampling error prevent the ensemble-estimated background error covariances from being optimal in the Kalman update equation (1). To address this, the background error covariances are inflated in amplitude (Anderson and Anderson 1999) and localized in space (Houtekamer and Mitchell 2001; Hamill et. al. 2001). Such ad-hoc measures are required to prevent a condition called filter divergence (Hamill et. al. 2001), in which the weight given to the ensemble first guess relative to the observations increases monotonically with time, while the first-guess drifts farther and farther away from the observations.

Both the CDAS-SFC and the Ensemble Filter use an NWP model to generate their first guess and dynamically propagate information forward to the next analysis time. The model is a 1998 version of the NCEP global medium range forecast model (Kanamitsu et. al. 1991). The model is spectral with a triangular truncation at wavenumber 62 (about 1.9° by 1.9° grid spacing) and has 28 sigma levels. The model physics are described in Wu and Iredell (1997). The model is the same as recently used to create a consistent series of re-forecasts from 1979 to the present (Hamill et al. 2004) and is still being used operationally at NCEP to generate certain probabilistic precipitation and near-surface temperature forecast products (NCEP 2005). Surface boundary conditions were taken from the NCEP-NCAR reanalysis and were the same for the CDAS-SFC and each ensemble member.

As a baseline for the CDAS-SFC and Ensemble Filter, we employ a statistical interpolation scheme with no dynamical model to propagate information to the next analysis. This scheme uses the 1971-2000 NCEP-NCAR reanalysis climatology as the first guess, and the climatological anomaly covariances as the background error covariances for that the first guess. Our Climatological Ensemble Filter (EnsClim) is implemented in a similar manner to the Ensemble Filter described above (eqs 1-4), but instead of using an ensemble of model forecasts integrated from the previous ensemble of analyses, a random ensemble of NCEP-NCAR reanalysis states from 1971-2000 is used as the first guess ensemble for each analysis. This system represents a Monte Carlo approximation to the climatological mean and climatological anomaly covariances.

Analyses created using only surface pressure observations

A representative example of a 500 hPa geopotential height analysis using the simulated 1895 network on 20 December 2001 0Z is shown in Fig. 2. Analyses produced with the three

data assimilation systems using only surface pressure observations are compared with the analysis from the full NCEP-NCAR reanalysis, which uses all available observations at all levels. The EnsClim analysis depicts many of the large-scale features associated with this time, including a substantial block over the North Atlantic and deep troughs over Europe and the North Pacific, but misses the smaller synoptic-scale features. In contrast, the CDAS-SFC analysis has many small-scale features, but they are positioned incorrectly, resulting in an error comparable to the EnsClim. The Ensemble Filter is able to represent not only the large-scale features, but also many of the synoptic-scale features and has an overall smaller error for this case.

Statistics summarizing the accuracy of the analyses are illustrated in Fig. 3 for geopotential height over the extratropical Northern Hemisphere. As an estimate of the analysis error and analysis quality, we have computed the root mean square difference (RMS) and anomaly pattern correlation with the full NCEP-NCAR reanalyses (Kalnay et al. 1996). For reference, we have compared the analysis error to the forecast errors of 1979-2001 re-forecasts made using this same model and the NCEP-NCAR reanalysis fields as initial conditions (Hamill et. al. 2004). We will refer to this forecast error as "modern NWP forecast error," as it is representative of operational NWP forecast errors over the period (Simmons and Hollingsworth, 2002)[2].

All three systems are capable of making 6-hourly height analyses that are "useful" as defined by errors smaller than the climatological standard deviation at every level in the

[2] Simmons and Hollingsworth (2002, their Figs. 6 and 8) show that operational ECMWF 500 hPa geopotential height RMS forecast error over the Northern Hemisphere varied from 1981 to 2001. As an example, the winter day 3 error was ~60 m in 1981, ~40 m in 1994, and ~30 m in 2001. The reforecast RMS error averaged over the period 1979-2001 for the same quantity (Fig. 3) is 47 m. Currently operational NWP forecasts, using much higher resolution models, have skill that is about one day better than our definition of modern NWP forecast error (A. Lorenc, 2005, personal communication). For example, the day 3 RMS error of the operational NCEP

troposphere. Each system produces average anomaly pattern correlations in the lower troposphere exceeding 0.7, a frequently used minimum for a useful correlation. The use of a forecast model appears to give the CDAS-SFC system no advantage over the EnsClim, suggesting either that the specified background error covariances for the CDAS-SFC system do not optimally represent the error of its forecasts or that the sparse observation network cannot constrain the evolution of a single model integration. We have ascertained that all three systems reproduce the large-scale aspects of the flow relatively well, but using the time-varying background error covariance estimated from a (relatively expensive) 100-member ensemble yields a substantial benefit in this data sparse case. Expected errors in the lower troposphere are comparable to the day 3 modern NWP forecast error, for example about 50 m at 500 hPa.

As more observations are added to the network, the relative superiority of the Ensemble Filter decreases in the lower troposphere, as illustrated for the 1905 experiment in Fig. 4. All three systems now well exceed pattern correlations of 0.7 for the lower troposphere, with correlations above 0.95 for many levels with the Ensemble Filter. The CDAS-SFC system produces a significantly better analysis than the EnsClim away from the surface. Results using the untuned CDAS (dotted curves) demonstrate the importance of tuning the 3DVAR system to account for the large forecast uncertainty when assimilating only surface observations.

In addition to having a relatively small error when averaged over the month, the analysis quality for each of the systems is fairly consistent at each time (Fig. 5). It may seem surprising that the CDAS-SFC surface system is not able to reconstruct the 500 hPa heights as well as the Ensemble Filter despite having the same specified boundary conditions and a model-propagated forecast as the first guess. Part of the reason may lie in the prescribed structure of the background error covariances which prevents a surface pressure observation from making a large

---

Global Forecast System for December 2004 was ~35 m (P. Caplan 2005 personal communication), comparable to the day 2 error in Fig. 3.

change to the upper-level first guess (See Sidebar Fig. SB2). In contrast, the Ensemble Filter can use a surface pressure observation to make a larger change in the first guess at upper-levels than at the surface, and a larger change away from the observation location than at that location. One can speculate that an advanced 3DVAR with reformulated, more complex background error covariances that vary spatially (e.g., Wu et. al. 2002, Purser et. al. 2003) or vary with the flow (e.g. Riishojgaard 1998, Buehner 2005) could result in a large improvement relative to the CDAS-SFC results.

Evidence that the time-varying nature of the background-error covariances in the Ensemble Filter system is contributing to its smaller error can be found in the analysis error time series (Fig. 5). The CDAS-SFC analysis error has large excursions from its mean value, reflecting those times when the static background–error covariances are a relatively bad (or good) estimate of the actual first guess error statistics. In contrast, the Ensemble Filter analysis error varies much less, suggesting that the flow-dependent background error covariances are able to produce analyses of more uniform quality.

The regional variation of analysis quality for the 1905 experiment is illustrated in Fig. 6 (left panels) for several variables using the Ensemble Filter 700 hPa analyses. The number of observations in each 2.5° gridbox is also shown. The variations of the 700 hPa wind and mass fields are both recovered well throughout the Northern Hemisphere extratropics despite the dramatically varying observational density used.

One sees a different story for the Tropics, where only some of the mass field variations are recovered with skill in regions of dense observations, such as over the tropical Atlantic and Indian Oceans. The wind field is poorly recovered, with correlations below 0.2 in many areas. Previous studies (e.g., Parrish and Derber 1992, Zagar et. al. 2005) have shown that a tropical pressure observation makes little change to the tropical wind first guess and, therefore, has little impact on the tropical wind analysis. In contrast, we find that the Ensemble Filter can produce a

substantial covariance between the first-guess surface pressure and tropical wind fields, resulting in modifications to the first-guess tropical wind field that are on the same order (0.5-1 m s^{-1}) as in the mid-latitudes. The resulting low quality of the Ensemble Filter tropical wind analyses (Fig. 6) suggests that the cross-covariances between tropical surface pressure and wind are in error. Whether it is indeed possible to produce useful tropical wind analyses using only surface pressure observations is an open research question.

Over the Southern Hemisphere, despite very sparse observations expected for 1905, Fig. 6 suggests that a surface pressure-based reanalysis should be able to reconstruct the lower tropospheric circulation of some regions at the start of the 20th century. Still, many regions, such as the South Pacific, are expected to be poorly analyzed, reflecting the lack of observations.

The results for the simulated 1935 network (Fig. 6 right panels) suggest that the analysis quality for the Northern Hemisphere extratropics will be consistent throughout the 20th century, while the quality for the Tropics and Southern Hemisphere will increase with additional observations. Correlations for the Tropical 700 hPa height field now are above 0.7 throughout most of the Tropics. However, the wind field remains poorly analyzed. For the Southern Hemisphere, the results are more encouraging. For instance, in the better-observed parts of the Southern Hemisphere extratropics (20°S-70°S) the areal fraction of 700 hPa zonal wind correlations above 0.7 goes from 29.6% for the 1905 network (Fig. 6b) to 51.3% for the 1935 network (Fig. 6f).

Overall, prospects appear very promising for high-quality analyses of the wintertime Northern Hemisphere circulation throughout the depth of the troposphere from the 1890's to present (Figs. 7, 8). Once a sufficient number and spatial coverage of surface pressure observations, spanning the hemisphere, is reached, additional observations do not further improve the upper-tropospheric analyses. In the middle and lower troposphere, the analyses continue to improve with additional observations. Comparing the results for the simulated 1895

network (red curves) to the 1905 network (blue curves) is particularly instructive. The 1905 network anticipates only 30 more observations per analysis, but these are all located in the North Pacific (e.g., Fig. 1). Additional observations for this region enhance the analysis quality and should be a priority in data recovery efforts.

The Northern Hemisphere 6-hourly meridional wind field is also reconstructed well throughout the depth of the troposphere (Fig. 8). These curves support the idea that synoptic variability could be reasonably well captured by reanalyses using only surface pressure observations. Meridional wind anomalies are important for quantities such as transports of heat, momentum, and moisture. More than 1 month of data will be necessary in future studies to compare the statistics of transports determined from analyses using only surface pressure with those from analyses using all observations. Still, the high skill in reconstructing the meridional wind, zonal wind (Fig. 6), and geopotential height (Figs. 6, 7) suggests that such second-order statistics may also be captured with reasonable accuracy.

Having demonstrated that consistent analyses for the 20th century could be possible for the Northern Hemisphere extratropics with networks similar to those currently digitized, we now further investigate other regions of the globe. Over the Tropics, the EnsClim, CDAS-SFC, and Ensemble Filter have similar analysis errors, providing some useful information about the mass field with the 1935 network (e.g., Fig. 6e, other methods not shown). It may be surprising that, for this region, the dynamically based methods have no advantage compared to the EnsClim in recovering even the 1000mb daily-averaged geopotential height field (Fig. 9). All three methods have pattern correlations that average just above 0.7, the lower limit of our definition of utility. The results show that the model propagation of the previous analysis field is not leading to a subsequent improvement in the analysis; a climatological first guess is performing just as well. Model errors (e.g., Hendon et. al. 2000) may be eliminating the expected advantage of the CDAS-SFC and Ensemble Filter systems over the EnsClim. Model improvement may be

necessary before more advanced data assimilation methods are able to improve upon the EnsClim. It is also possible that an empirically derived dynamical model may outperform a numerical one for this purpose, in the same way that an empirical model outperforms the tropical forecasts of this numerical model beyond a week (Newman et. al., 2003). Increasing the number of surface pressure observations may help the analysis, but even the full 2001 network produces anomaly pattern correlations for geopotential height that are still just above 0.8 at 1000 hPa and fall below 0.7 above 700 hPa (not shown).

For the Southern Hemisphere extratropics, in contrast to the Tropical results, increasing the number of observations will dramatically improve the analysis quality (Fig. 10). A useful analysis appears possible after 1935. This result is consistent for other variables (not shown). It is interesting to note that the analysis quality *decreases* from 1895 to 1915. This behavior arises from the decreasing number of pressure observations over this period in the current digital archive (not shown). Recovering additional pressure observations for the Southern Hemisphere can be expected to increase the analysis quality throughout the troposphere commensurate with the observational network (Fig. 6).

It is also important to investigate how the analysis quality will vary through the seasonal cycle. Results for analyses of June 2001 indicate that upper tropospheric analyses in the Northern Hemisphere summer will not be as good as during winter but will still have useful information (Fig. 11). Experiments were performed for June 2001 using only 1905 and 2001 networks and the Ensemble Filter to suggest the range of error expected as the network varies. In the upper troposphere, analysis errors are comparable to the day 4 modern NWP forecast error, and anomaly pattern correlations are on the order of 0.8. In the lower troposphere, additional observations will improve the analysis quality, with errors for the 1905 network comparable to the 2 to 3 day forecast error. There is a strong similarity between the summer results of the two hemispheres, December in the Southern Hemisphere (Fig. 10) and June in the

Northern Hemisphere (Fig. 11). In the Southern Hemisphere for June 2001, results improve somewhat compared to December 2001, with geopotential height pattern correlations using the 2001 network greater than 0.85 in the upper-troposphere (not shown).

Conclusions

The original goal of this study was to determine whether modern data assimilation systems could use the available historical surface pressure observations to produce accurate daily surface weather maps for the period of the 19th century to the present. We have concluded that, not only are surface weather maps feasible, but the surface pressure observations could be used to reanalyze the entire extratropical tropospheric circulation.

For the Northern Hemisphere winter, such a reanalysis using the Ensemble Filter can be expected to be of high quality, as measured by analysis errors that are substantially less than the climatological standard deviation for all dynamical variables. We expect the analysis error to be on the order of the modern NWP 1 to 2 day forecast error in the lower troposphere and the 2 to 3 day forecast error in the middle and upper troposphere. This modern NWP forecast error is based on 1979-2001 re-forecast skill using a 1998 NCEP model and NCEP-NCAR reanalysis fields as initial conditions. Currently operational NWP forecasts have forecast errors about one day better than shown here (A. Lorenc, 2005 personal communication; P. Caplan, 2005 personal communication).

Summertime analysis errors for the Northern Hemisphere will be relatively larger than in winter (closer to the modern NWP 3-day forecast error in the lower troposphere and the 4-day error in the upper troposphere) but still much less than the climatological standard deviation throughout the depth of the troposphere. In contrast, in the Tropics and the Southern Hemisphere, currently available surface pressure observations are insufficient for a useful

reanalysis before 1935. Additional digitized observations will aid in producing a better analysis over the globe, with pressure observations over the Pacific Ocean being the highest priority.

One previously noted problem with the current generation of reanalyses is their inhomogeneity arising from the time-changing observational network. While there is no subset of station and marine pressure observations that could be kept constant to avoid all inhomogeneities, the results here suggest that a surface-pressure based reanalysis could be more consistent than previous reanalyses, particularly in the upper troposphere, as the observational network varies. Further inhomogeneities may be introduced if the observational error changes over the period of record. Recent evidence suggests that the measurement error characteristics of specific collections within the surface pressure station and marine observational database may vary over time (e.g. Kent and Berry 2005, Ishii et. al. 2005, Slonosky and Graham 2005), but the general error characteristics of the complete database are not known. The results of Velicogna et. al. (2001) suggest that half to three-quarters of the specified observational error in station pressure observations is assigned based on the error in representing a large area (i.e., a model grid box) by a point measurement. This component of the observational error, the so-called "error of representativeness," will largely be a function of the model resolution and not the measurement accuracy. Therefore, the overall observational error may vary less over the historical period than one might suppose. As estimates of historical observational errors are produced they can be incorporated in the analysis (i.e., eqs 3-4). Since the Ensemble Filter produces estimates of analysis error (eq 2), the effect of inhomogeneities in observation networks and observation errors can easily be quantified.

Additional improvements to the results presented here may be possible. Hourly or even more frequent assimilation of surface pressure observations may improve the upper-level analysis, as has been demonstrated with a linear quasi-geostrophic system (Bengtsson 1980), with a real 4DVAR application (Jarvinen et. al. 1999), and with a perfect model ensemble filter

experiment (Anderson et. al. 2005). Using observations collected past the analysis time may also improve the analysis quality. Unlike an operational analysis, a reanalysis is not constrained to use observations that only precede the analysis time. An ensemble Kalman smoother algorithm, which generalizes the Kalman filter to use observations both before and after the analysis time, has been proposed (Whitaker and Compo 2002) for this purpose, building on the work of Cohn et al. (1994). We will investigate whether improvements to the results presented here are possible using frequent or "future" observations.

Having demonstrated that reanalyses using only surface pressure observations are feasible with an expected accuracy much less than climatological error, we have left unresolved the issue of whether such reanalyses will be useful for estimating decadal and longer term trends in the statistics of weather and climate, and especially in the statistics of extreme phenomena. While the results presented here are compelling, they are not conclusive in this regard. Additional observing system experiments could answer these questions more definitively.

Sidebar 1: Available Historical Surface and Sea Level Pressure Observations

Comparing the quantities of surface pressure observations available now to that available even 10 years ago illustrates why an historical reanalysis should be considered now (Fig. SB1). We have estimated the number of digitized surface and sea level pressure observations in 5°X5° squares around the globe in each month from 1854 to 1955. Our estimate includes unique land observations from the National Climatic Data Center (NCDC) and National Center for Atmospheric Research (NCAR) data archives and marine observations from the Comprehensive Ocean Atmosphere Dataset (COADS) Release 2 (Woodruff et al. 1998). Over the Northern Hemisphere extratropics, the data available from these sources (red curve) have been greatly enhanced by several observation recovery efforts (blue curve). Many more marine pressure data

are now available in the International Comprehensive Ocean Atmosphere Dataset (ICOADS 2.1; Woodruff et. al. 2005; Worley et. al. 2005) compared to the previous COADS dataset. Some digitized marine datasets have not yet been fully blended into ICOADS, such as the Kobe Collection 2001 (Manabe, 1999) and German Marine Meteorological archive (V. Wagner, personal communication 2003). The unique additional observations from these collections are included in the blue curve.

The pink curve in Fig. SB1 includes "proxy" data and current digitization efforts. Prior to 1912, the US Merchant Marine collection has not yet been digitized (J. Elms, personal communication 2005), but the Northern Hemisphere SLP Grids based on the USHWM contain these data in the analyzed maps (United States Weather Bureau 1944, e.g., Fig. 1). We have included these grids from 20°N-70°N at their digitized 5°X10° resolution only over the oceans in our estimate, assuming 70% of those points would be usable. The recently digitized United States Marine Meteorological Journals (S. Woodruff, personal communication 2004) are now available and included in the pink curve. Land data recovery efforts, such as from the NOAA Climate Data Modernization Program (CDMP, J. Elms personal communication 2005), European Union projects (Ansell et. al. 2005, submitted to J. Climate), the Waves and Storms dataset (Schmith et. al. 1997, The WASA Group 1998), Environment Canada pressure station dataset (V. Swail, personal communication 2005), and other international efforts (R. Jenne, personal communication 2004) further raise the number of observations expected to be available digitally by the end of 2005. We have attempted to simulate surface pressure networks for these data densities.

We have also made a conservative estimate of the number of additional observations that could be available by assuming that the Global Climate Historical Network (GHCN, Vose et. al. 1992) monthly station pressure observations could be available once per day (Fig. SB1 black curve). This is conservative in the sense that many stations reported two or more times per day,

and the GHCN does not include all reporting stations. Further, we have made no attempt to estimate the number of undigitized marine observations, which for the Kobe Collection alone could be more than 6.5 million global reports from 1890-1930 (Manabe 1999). Volumes of data undigitized from sources such as UK logbooks, the US Merchant Marine Collection prior to 1912, and the US Navy are unknown but estimated to be in the tens of millions of reports for this period (S. Woodruff, personal communication 2005).

Sidebar 2: Influence of a single pressure observation on the resulting analysis.

To compare the effect of pressure observations in the Ensemble Filter and CDAS-SFC systems, we have conducted four single observation assimilation experiments. Two "observations" were assimilated separately by both systems. Each observation was prescribed to have a value 1 hPa larger than the surface pressure forecast for 25 December 2001 06Z from the previous assimilation using a 1905 network. The results of the two separate experiments are plotted together in the Figure to make a summary of the results; it is evident that neither observation would significantly influence the analysis of the other.

The filled contours show the first guess geopotential height field at 1000 hPa (bottom) and 300 hPa (top) from the CDAS-SFC (left) and Ensemble Filter (right). The line contours show the analysis increment, the difference between the analysis after assimilating the indicated observation and the first guess field. For the CDAS-SFC system the analysis increment of a single observation does not depend on longitude. For surface pressure observations, the increment is centered on the observation location, is largest at the surface, and decreases with the height. Increments are a maximum of 6.9 m for 1000 hPa heights and 5.5 m for 300 hPa heights.

The right panels illustrate the ability of the Ensemble Filter to create spatially inhomogeneous background error covariances that change with the flow and observational density. The analysis increments produced by the two observations are quite different, commensurate with the larger expected uncertainty in the observation-poor central Pacific and

smaller uncertainty in the observation-rich continental North America. The larger uncertainty becomes a larger analysis increment. The uncertainty in the first guess, the background error covariance, over the mid-Pacific translates an observation 1 hPa above the background into a general weakening of the nearby trough, with a maximum weakening at 1000 hPa of 8.8 m directly to the *east* of the observation. The background error covariance further translates the single observation into a weakening of the upper-level trough, but with maximum weakening at 300 hPa of 10.5 m to the *southeast* of the observation location. In this case, a single surface pressure observation is producing an analysis increment that is tilted with height and having maximum amplitude in the upper troposphere. The Ensemble Filter also changes the sign of the increment to the northwest of the observation. Even over the interior continent, the effect of the single observation, though smaller, still has maximum amplitude in the upper troposphere and is tilted with height. Most likely, the ability to produce analysis increments that vary with the flow and the observation network gives the Ensemble Filter its advantage in the extratropics. A more complex 3DVAR formulation could also possess these flow-dependent properties (e.g., Riishojgaard 1998, Buehner 2005) and may have a similar advantage over the CDAS-SFC system.

Acknowledgments: Useful discussions with J. Anderson, J. Barsugli, R. Dole, T. Hamill, K. Hodges, E. Kalnay, A. Simmons, J.-N. Thepaut, and H. van den Dool are gratefully acknowledged. R. Allan, T. Ansell, S. Doty, J. Elms, R. Jenne, P. Jones, S. Lubker, T. Schmith, V. Swail, V. Wagner, and S. Woodruff provided invaluable assistance in determining current and future data availability. The NOAA-CIRES Climate Diagnostics Center IT staff provided invaluable computer support. Department of Commerce/Boulder and NCAR librarians kindly assisted with manuscript research. The comments of two anonymous reviewers and A. Lorenc greatly improvement the clarity of the manuscript. The NOAA Forecast System Laboratory

graciously provided computer time on the High Performance Computing System. This research was conducted with the support of an Innovative Research Program grant from the University of Colorado's Cooperative Institute for Research in the Environmental Sciences, a grant from NOAA's Office of Atmospheric Research, and a grant from NOAA's Office of Global Programs.

References:

Anderson, J. L., and S.L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.,* **127**, 2741–2758.

_____, B. Wyman, S. Zhang, and T. Hoar, 2005: Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *J. Atmos. Sci.*, **62**, 2925-2938..

Ansell, T., and Coauthors, 2005: Daily mean sea level pressure reconstructions for the European - North Atlantic region for the period 1850-2003. J. Climate, submitted.

Arkin, P., E. Kalnay, J. Laver, S. Schubert, and K. Trenberth, 2003: Ongoing Analysis of the Climate System: A Workshop Report, 18-20 August 2003, Boulder, Colorado.

[cited 2005, available online at http://www.ofps.ucar.edu/joss_psg/meetings/climatesystem/FinalWorkshopReport.pdf]

Basist, A.N., and M. Chelliah, 1997: Comparison of tropospheric temperatures derived from NCEP/NCAR reanalysis, NCEP operational analysis, and the Microwave Sounding Unit. *Bull. Am. Met. Soc.*, **78**, 1431-1447.

Bengtsson, L., 1980: On the use of a time sequence of surface pressures in four-dimensional data assimilation. *Tellus*, *32*, 189-197.

_____, and J. Shukla, 1988: Integration of Space and In Situ Observations to Study Global Climate Change. *Bull. Amer. Met. Soc.,* **69**, 1130–1143.

_____, S. Hagemann and K. I. Hodges, 2004a: Can climate trends be calculated from reanalysis data? *J. Geophys. Res.*, **109**, D11111, doi:10.1029/2004JD004536.

_____, K. I. Hodges and S. Hagemann, 2004b: Sensitivity of the ERA40 reanalysis to the observing system: determination of the global atmospheric circulation from reduced observations. *Tellus*, **56A**, 456-471.

Bronnimann, S., G.P. Compo, P.D. Sardeshmukh, R. Jenne, and A. Sterin, 2005: New approaches for extending the Twentieth Century climate record. *Eos*, **86**, 2-6.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP Setting. *Q. J. Roy. Met. Soc.,* **131**, 1013-1043.

Burgers, G., P.J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126**, 1719–1724.

Chang, E.K.M., and Y. Fu, 2003: Using mean flow change as a proxy to infer interdecadal storm track variability. *J. Climate*, **16***,* 2178-2196.

Cohn, S.E., N.S. Sivakumaran, and R. Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.,* **122**, 2838–2867.

Evensen, G. 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143-10 162.

Group on Earth Observations, 2005: The Global Earth Observation System of Systems 10-year Implementation Plan, cited 2005:. [Available online at http://earthobservations.org/docs/10-Year Implementation Plan.pdf.]

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter- 3D-variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

____, J.S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.,* **129**, 2776–2790.

____, ____, and X. Wei, 2004: Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.,*

Harnik, N. and E.K.M. Chang, 2003: Storm track variations as seen in radiosonde observations and reanalysis data. *J. Climate*, **16**, 480-495.

Hayden, B.P., 1984: A systematic error in the Northern Hemisphere sea-level pressure data set. *Mon. Wea. Rev.,* **112**, 2354–2356.

Hendon, H., B. Liebmann, M. Newman, J.D. Glick, and J.E. Schemm, 2000: Medium range forecast errors associated with active episodes of the Madden-Julian oscillation. *Mon. Wea. Rev.*, **128**, 68-86.

Hodges, K.I., B. J. Hoskins, J. Boyle and C. Thorncroft, 2003: A Comparison of Recent Reanalysis Datasets Using Objective Feature Tracking: Storm Tracks and Tropical Easterly Waves. *Mon. Wea. Rev.*, **131**, 2012-2037.

Houtekamer, P.L. and H.L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.,* **129**, 123–137.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analysis of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865-879.

Jarvinen, H., E. Andersson, and F. Bouttier, 1999: Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus*, **51A**, 469-488.

Jones, P. D., 1987: The early twentieth century Arctic high—Fact or fiction? *Climate Dyn.,* **1**, 63–75.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Met. Soc.*, **77***,* 437-471.

Kanamitsu, M., and S.-O. Hwang, 2005: The role of sea surface temperature in reanalysis. *Mon. Wea. Rev.*, in press.

_____, and Coauthors 1991: Recent changes implemented into the global forecast system at NMC. *Wea. Forecasting*, 6, 425-435.

Kaplan, A., Y. Kushnir, and M.A. Cane, 2000: Reduced space optimal interpolation of historical marine sea level pressure: 1854-1992, *J. Climate*, **13**, 2987-3002.

_____, M.A. Cane, Y. Kushnir, A.C. Clement, M.B. Blumenthal, B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856-1991. *J. Geophys. Res.,* **103**, 18,567-18,589.

Kent, E.C., and D.I. Berry, 2005: Quantifying random measurement errors in voluntary observing ships' meteorological observations. *Int. J. Climatol.*, **25**, 843-856.

Kinter, J.L, M. J. Fennessy, V. Krishnamurthy and L. Marx. 2004: An Evaluation of the Apparent Interdecadal Shift in the Tropical Divergent Circulation in the NCEP-NCAR Reanalysis. *J. of Climate*, **17**, 349-361.

Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bull. Am. Met. Soc.,* **82**, 247-268.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112**, 1177–1194.

Mahoney and Coauthors, 2003: Strategic Plan for the U.S. Climate Change Science Program. Report by the Climate Change Science Program and the Subcommittee on Global Change Research [available online at http://climatescience.gov/Library/stratplan2003/default.htm]

Manabe, T., 1999: The digitized Kobe collection, phase I: Historical surface marine meteorological observations in the archive of the Japan Meteorological Agency. *Bull. Amer. Meteor. Soc.,* **80**, 2703–2715.

Mason, P.J., and Coauthors, 2004: Implementation plan for the global observing system for climate in support of the UNFCCC. WMO-IOC-UNEP-ICS, GCOS-92, Technical Document No. 1219, World Meteorological Organization, Geneva, Switzerland, 153pp.

NCAR, cited 2005: Daily Northern Hemisphere Sea Level Pressure Grids, continuing from 1899, Data Sources. [available online at http://dss.ucar.edu/datasets/ds010.0/docs/]

NCEP, cited 2005: CDC MRF Reforecast [available online at http://www.cpc.ncep.noaa.gov/products/predictions/short_range/tools/cdc.pub.html]

Newman M., P. D. Sardeshmukh, and J. W. Bergman, 2000: An assessment of the NCEP, NASA and ECMWF reanalyses over the Tropical West Pacific warm pool. *Bull. Amer. Meteor. Soc.*, **81**, 41-48.

_____, _____, C. R. Winkler, and J. S. Whitaker, 2003: A study of subseasonal predictability. *Mon. Wea. Rev.*, **131**, 1715-1732.

Parrish, D.F., and J.C. Derber, 1992: The National Meteorological Center's spectral statistical–interpolation analysis system. *Mon. Wea. Rev.,* **120**, 1967–1988.

Purser, R.J., W.-S. Wu, D.F. Parrish, and N.M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.,* **131**, 1536–1548.

Rayner, N.A., D.E. Parker, E.B. Horton, C.K. Folland, L.V. Alexander, D.P. Rowell, E.C Kent, and A. Kaplan, 2003: Global analyses of SST, sea ice and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.

Riishojgaard, L.P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus*, **50A**, 42-57.

Simmons, A.J., and J.K. Gibson, 2000: *The ERA-40 Project Plan*, ERA-40 Project Report Series No. 1 ECMWF. Shinfield Park, Reading, UK, 63pp.

_____, and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Met. Soc.*, **128**, 647-677.

Schmith, T., H. Alexandersson, and H. Tuomenvirta, 1997: North Atlantic–European pressure observations 1868–1995 (WASA dataset version 1.0). Danish Meteorological Institute Tech. Rep. 97-3, 13 pp.

Slonosky, V.C., and E. Graham, 2005: Canadian pressure observations and circulation variability: Links to air temperature. *Int. J. Climatol.,* doi: 10.1002/joc.1191.

Smith, T.M, and R. W. Reynolds, 2004a: Improved Extended Reconstruction of SST (1854-1997). *J. Climate*, **17**, 2466-2477.

____, and ____, 2004b: Reconstruction of Monthly Mean Oceanic Sea Level Pressure Based on COADS and Station Data (1854-1997). *J. Atmos. and Oceanic Tech*., **21**, 1272-1282.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill, and J.S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.,* **131**, 1485–1490.

Trenberth, K. E., and D.A. Paolino, 1980: The Northern Hemisphere sea-level pressure dataset: Trends, errors and discontinuities. *Mon. Wea. Rev.,* **108**, 855–872.

_____, and L. Smith, 2005: The Mass of the Atmosphere: A Constraint on Global Analyses. *J. Climate*, **18**, 864-875.

United States Weather Bureau, 1944: *Daily Synoptic Series, Historical Weather Maps, Northern Hemisphere Sea-Level, January 1899 to June 1939*. Cooperative project of U.S. Army, Air Force, and Weather Bureau Statistics Division.

Velicogna, I., J. Wahr, H. Van den Dool, 2001: Can surface pressure be used to remove atmospheric contributions from GRACE data with sufficient accuracy to recover hydrological signals? *J. Geophys. Res.*, **106**, 16415-16434, 10.1029/2001JB900228, 2001.

Vose, R. S., R. L. Schmoyer, P. M. Steurer, T. C. Peterson, R. Heim, T. R. Karl, and J. K. Eischeid, 1992: The global historical climatology network: Longterm monthly temperature, precipitation, sea level pressure, and station pressure data. Technical Report NDP-041, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

The WASA Group, 1998: Changing waves and storms in the Northeast Atlantic? *Bull. Amer. Meteor. Soc.*, **79**, 741-760.

Whitaker, J. S., and G.P. Compo, 2002: An ensemble Kalman smoother for reanalysis. *Proc. Symp. on Observations, Data Assimilation and Probabilistic Prediction,* Orlando, FL, Amer. Meteor. Soc., 144–147.

_____, and T.M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.,* **130**, 1913–1924.

_____, G. Compo, X. Wei, T. Hamill, and P. Sardeshmukh, 2003: Feasibility of R1895 using only surface pressure observations. Workshop on Ongoing Analysis of the Climate System, Boulder, CO 18-20 August 2003, [available online at http://www.ofps.ucar.edu/joss_psg/meetings/climatesystem/]

_____, G.P. Compo, X. Wei and T.M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.,* **132**, 1190–1200.

Williams, J. and H. van Loon, 1976: An examination of the Northern Hemisphere sea-level pressure data set. *Mon. Wea. Rev.,* **104**, 1354–1361.

Woodruff, S., H. Diaz, J. Elms, and S. Worley, 1998: COADS release 2 data and metadata enhancements for improvements of marine surface flux fields. *Phys. Chem. Earth.,* **23**, 517–527.

____, ____, S.J. Worley, R.W. Reynolds, and S.J. Lubker, 2005: Early ship observational data and ICOADS, Climatic Change, in press.

Worley, S.J., S.D. Woodruff, R.W. Reynolds, S.J. Lubker, and N. Lott, 2005: ICOADS Release 2.1 data and products, *Int. J. Climatol.*, **25**, 823-842.

Wu, W., and Iredell M., 1997: Changes to the 1997 NCEP operational MRF model analysis/forecast system. NCEP Technical Procedures Bulletin 443, National Weather Service, Office of Meteorology, Programs and Plans Division, Silver Spring, MD. [Available online at www.nws.noaa.gov/om/tpb/indexb.htm.].

Wu, W.-S., R.J. Purser, and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.,* **130**, 2905-2916.

Zagar, N., E. Anderson, and M. Fisher, 2005: Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Q. J. Roy. Meteor. Soc.*, **131**, 987-1011.

Figure Captions

Figure 1. United States Historical Weather Map of Sea Level Pressure for 1 February 1899 1300 GMT (United States Weather Bureau 1944). Individual observations used in the map are indicated with synoptic symbols. For this map, 27 marine reports from the US Merchant Marine Collection are present in the Pacific Ocean. The hemispheric network is comparable to what we have used for our 1905 assimilation experiments.

Figure 2. Comparison of analyses of 20 December 2001 00Z 500 hPa geopotential height from (top left) full NCEP-NCAR reanalysis using all available observations at all levels (> 150,000) and parallel assimilation experiments with a simulated 1895 network of only 308 surface pressure observations from (top right) Climatological Ensemble Filter (EnsClim) [RMS difference with full NCEP-NCAR reanalysis is 95.7 m], (bottom left) Ensemble Filter (EnsFilt) [RMS difference with full reanalysis is 49.2 m], and (bottom right) CDAS-SFC [RMS difference with full reanalysis is 96.0 m]. Blue dots indicate the location of the surface pressure observations used to make the experimental analyses. The 5500 m line is thickened, and the contour interval is 50 m.

Figure 3. Variation of Northern Hemisphere skill analyzing December 2001 using different assimilation systems and the same surface pressure observations simulating an 1895 network. (left) Colored curves show the monthly-averaged 6-hourly root-mean-square (RMS) differences over the Northern Hemisphere (20N-90N) between geopotential height analyses from full NCEP-NCAR reanalysis and assimilation experiments using only surface pressure observations at 1895 densities and Climatological Ensemble Filter (EnsClim)), CDAS-SFC, and Ensemble Filter (EnsFilt). Thin solid black curves show RMS error of *modern* NWP forecasts at days 1 through 5. Thick solid black curve shows climatological standard deviation. (right) Colored

curves show the monthly average of 6-hourly anomaly pattern correlations over the Northern Hemisphere of geopotential height anomalies from full NCEP-NCAR reanalyses and assimilation experiments.

Figure 4: As in Fig. 3 but for simulated 1905 observational densities. Results using the untuned CDAS are indicated by the dotted curves.

Figure 5. Time series of the 500 hPa geopotential height root mean square difference between analyses of December 2001 from the full NCEP-NCAR reanalysis and from the assimilation experiments. The RMS is averaged over the Northern Hemisphere (20°N-90°N). Climatological standard deviation is indicated by the thick black line.

Figure 6. Local anomaly correlation of December 2001 4-times daily analyses from the full NCEP-NCAR reanalysis and (left) 1905 and (right) 1935 assimilation experiments using the Ensemble Filter. Correlations are shown for (a),(e) 700 mb geopotential height; (b),(f) 700 mb zonal wind; and (c),(g) 700 mb meridional wind. Colors in the bottom panels indicate the number of surface pressure observations used in each 2.5° X 2.5° gridbox.

Figure 7. Variation of skill analyzing December 2001 using different simulated networks and the Ensemble Filter. (left) Colored curves show the average root-mean-square (RMS) difference over the Northern Hemisphere (20°N-90°N) between 6-hourly geopotential height analyses from full NCEP-NCAR reanalysis and Ensemble Filter assimilation experiments using only surface pressure observations at 1895, 1905, 1915, 1935, and 2001 densities. Thin solid black curves show *modern* RMS error of NWP forecasts at days 1 through 5. Thick solid black curve shows

climatological standard deviation. (right) Colored curves show the monthly average of 6-hourly anomaly pattern correlations over the Northern Hemisphere of geopotential height anomalies from full NCEP-NCAR reanalyses and assimilation experiments.

Figure 8. Variation of skill over the Northern Hemisphere analyzing December 2001 meridional wind using varying simulated networks and the Ensemble Filter. As in Fig. 7.

Figure 9. Time series of daily-averaged 1000 hPa height anomaly pattern correlation over the Tropics (20°N-20°S). Assimilation experiment analyses of December 2001 using the simulated 1935 surface pressure network are correlated with the full NCEP-NCAR reanalyses.

Figure 10. As in Fig. 7 for December 2001 but over the Southern Hemisphere (20°S-90°S).

Figure 11. As in Fig. 7 but skill for June 2001 over the Northern Hemisphere using only simulated 1905 and 2001 networks and Ensemble Filter.

Figure Sidebar 1. Our estimate of the number of Northern Hemisphere (20°N-90°N) surface pressure observations available per day from land and marine sources. Red curve: based on previous COADS and land sources. Blue curve: now available including ICOADS Release 2.1 and Kobe Collection 2000. Pink curve: includes digitization underway, US Marine Meteorological Journals, and quality control of the US Historical Weather Map Series. Black curve: includes international and US data available in manuscript form in NCDC archives. Years indicated with "Ex" show the number of observations used per day in the assimilation experiments.

Figure Sidebar 2. Effect of a single surface pressure observation (gray dot) on the final analysis. Filled contours show the first guess field from (a,c) CDAS-SFC and (b,d) Ensemble Filter for geopotential height at (a,b) 300 hPa and (c,d) 1000 hPa. Line contours show the analysis increment after assimilating a single pressure observation that is 1 hPa larger than the first guess pressure field at the indicated location (gray dots). The two gray dots represent two separate experiments. First guess (filled) contour interval is 100 hPa. Analysis increment (line) contour interval is 2 m starting at 1m. Positive (negative) increments are indicated with black (red) contours.

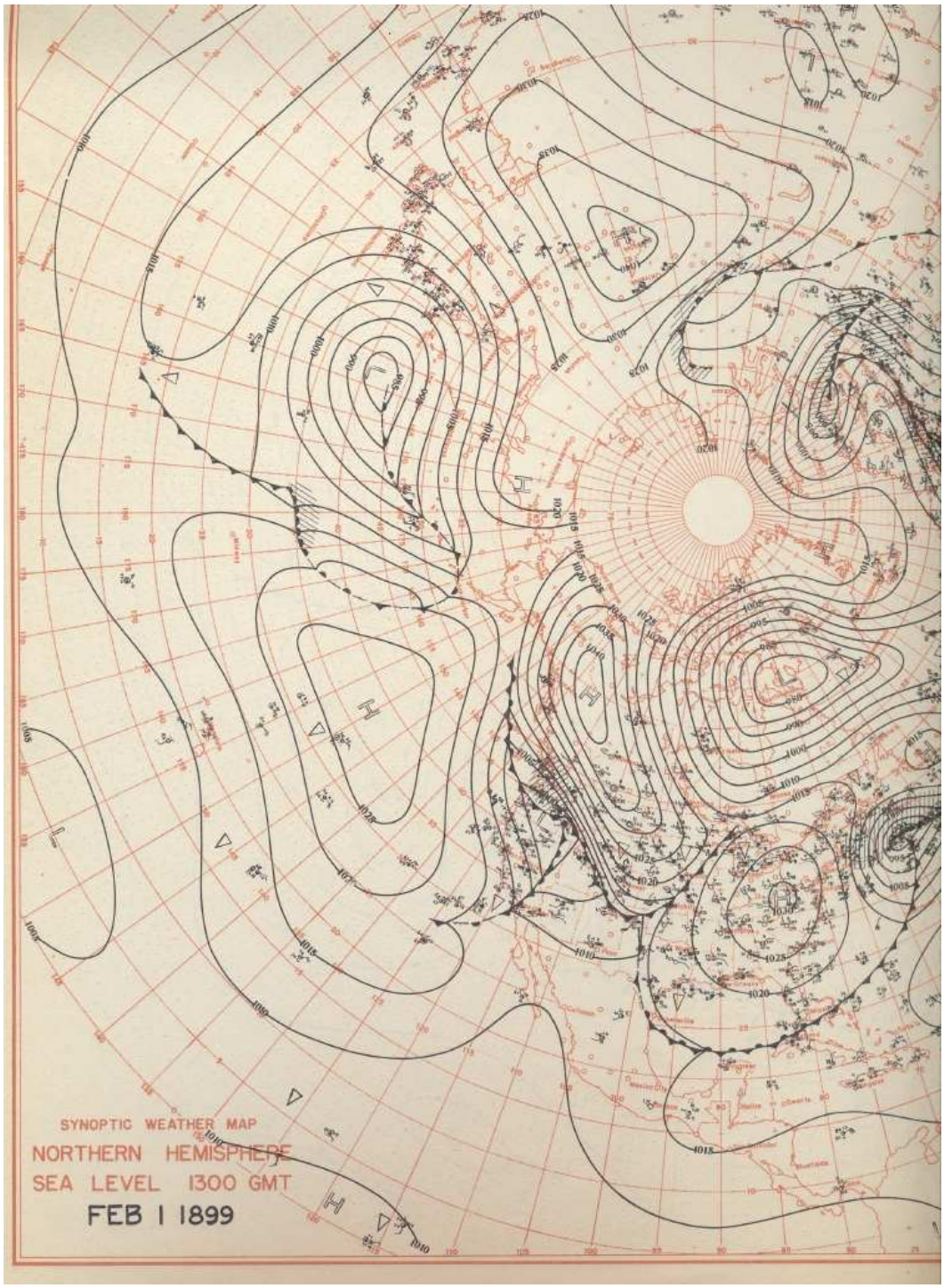

Figure 1. United States Historical Weather Map of Sea Level Pressure for 1 February 1899 1300 GMT (United States Weather Bureau 1944). Individual observations used in the map are indicated with synoptic symbols. For this map, 27 marine reports from the US Merchant Marine Collection are present in the Pacific Ocean. The hemispheric network is comparable to what we have used for our 1905 assimilation experiments.

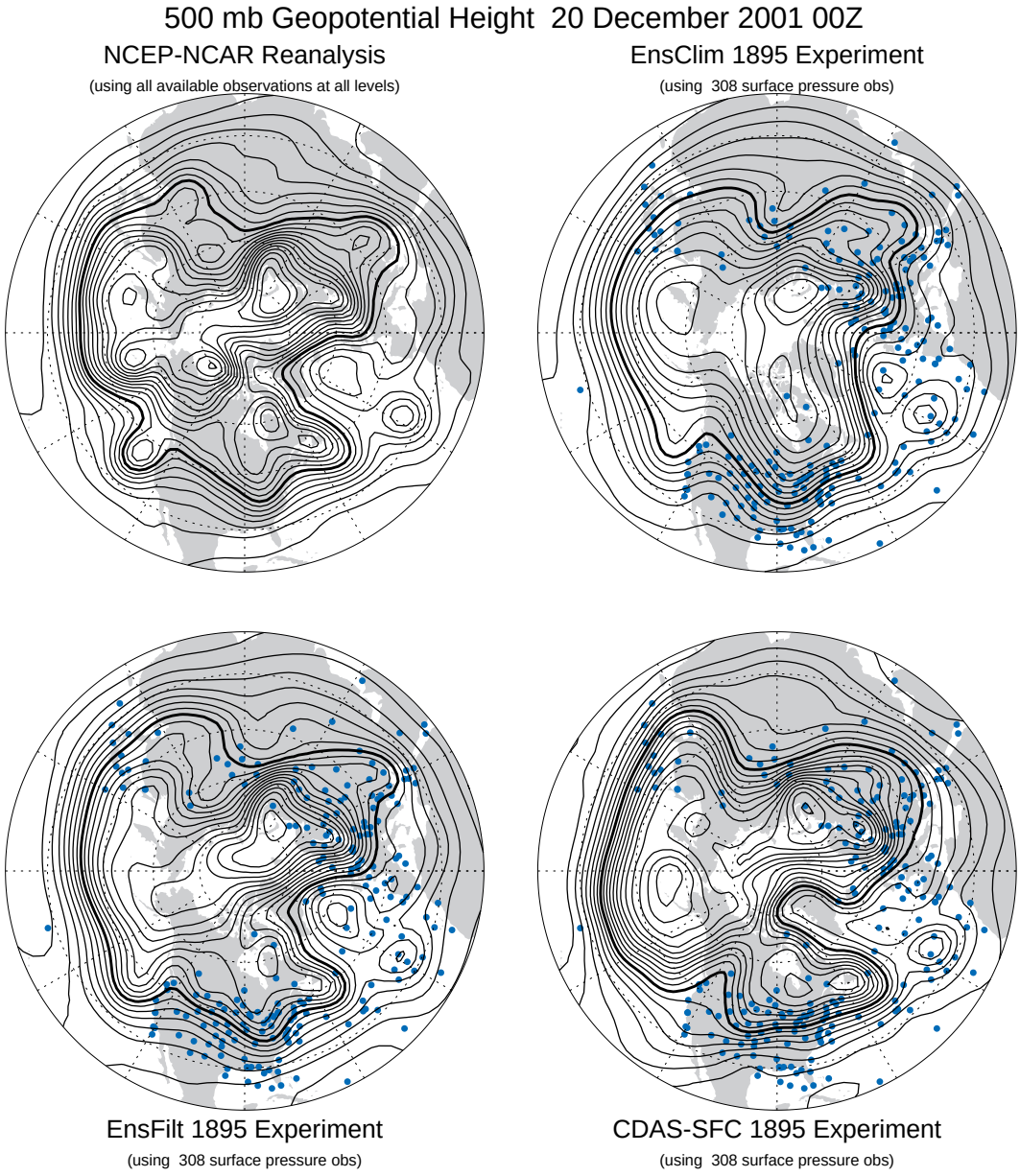

Figure 2. Comparison of analyses of 20 December 2001 00Z 500 hPa geopotential height from (top left) full NCEP-NCAR reanalysis using all available observations at all levels (> 150,000) and parallel assimilation experiments with a simulated 1895 network of only 308 surface pressure observations from (top right) Climatological Ensemble Filter (EnsClim) [RMS difference with full NCEP-NCAR reanalysis is 95.7 m], (bottom left) Ensemble Filter (EnsFilt) [RMS difference with full reanalysis is 49.2 m], and (bottom right) CDAS-SFC [RMS difference with full reanalysis is 96.0 m]. Blue dots indicate the location of the surface pressure observations used to make the experimental analyses. The 5500 m line is thickened, and the contour interval is 50 m.

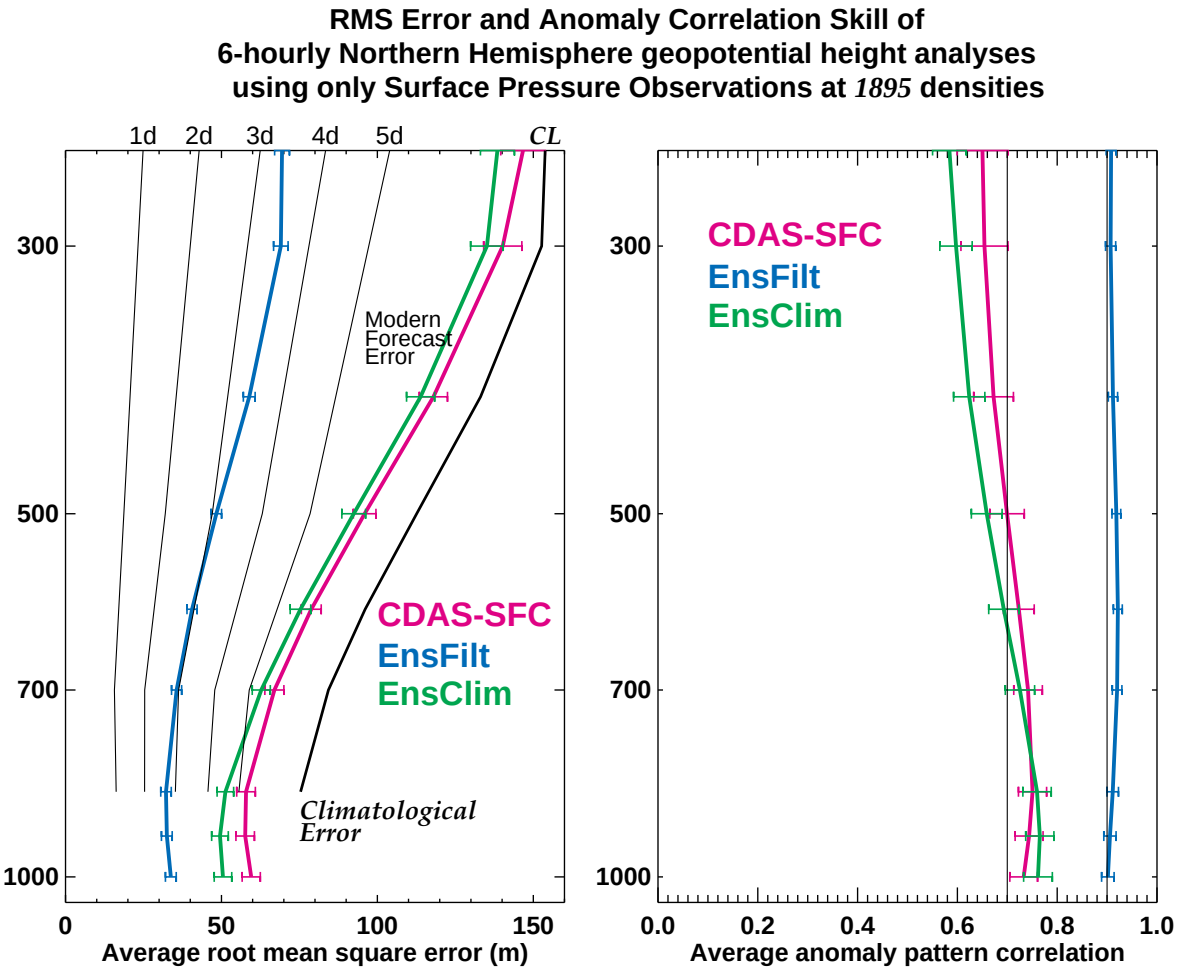

Figure 3. Variation of Northern Hemisphere skill analyzing December 2001 using different assimilation systems and the same surface pressure observations simulating an 1895 network. (left) Colored curves show the monthly-averaged 6-hourly root-mean-square (RMS) differences over the Northern Hemisphere (20N-90N) between geopotential height analyses from full NCEP-NCAR reanalysis and assimilation experiments using only surface pressure observations at 1895 densities and Climatological Ensemble Filter (EnsClim), CDAS-SFC, and Ensemble Filter (EnsFilt). Thin solid black curves show RMS error of *modern* NWP forecasts at days 1 through 5. Thick solid black curve shows climatological standard deviation. (right) Colored curves show the monthly average of 6-hourly anomaly pattern correlations over the Northern Hemisphere of geopotential height anomalies from full NCEP-NCAR reanalyses and assimilation experiments.

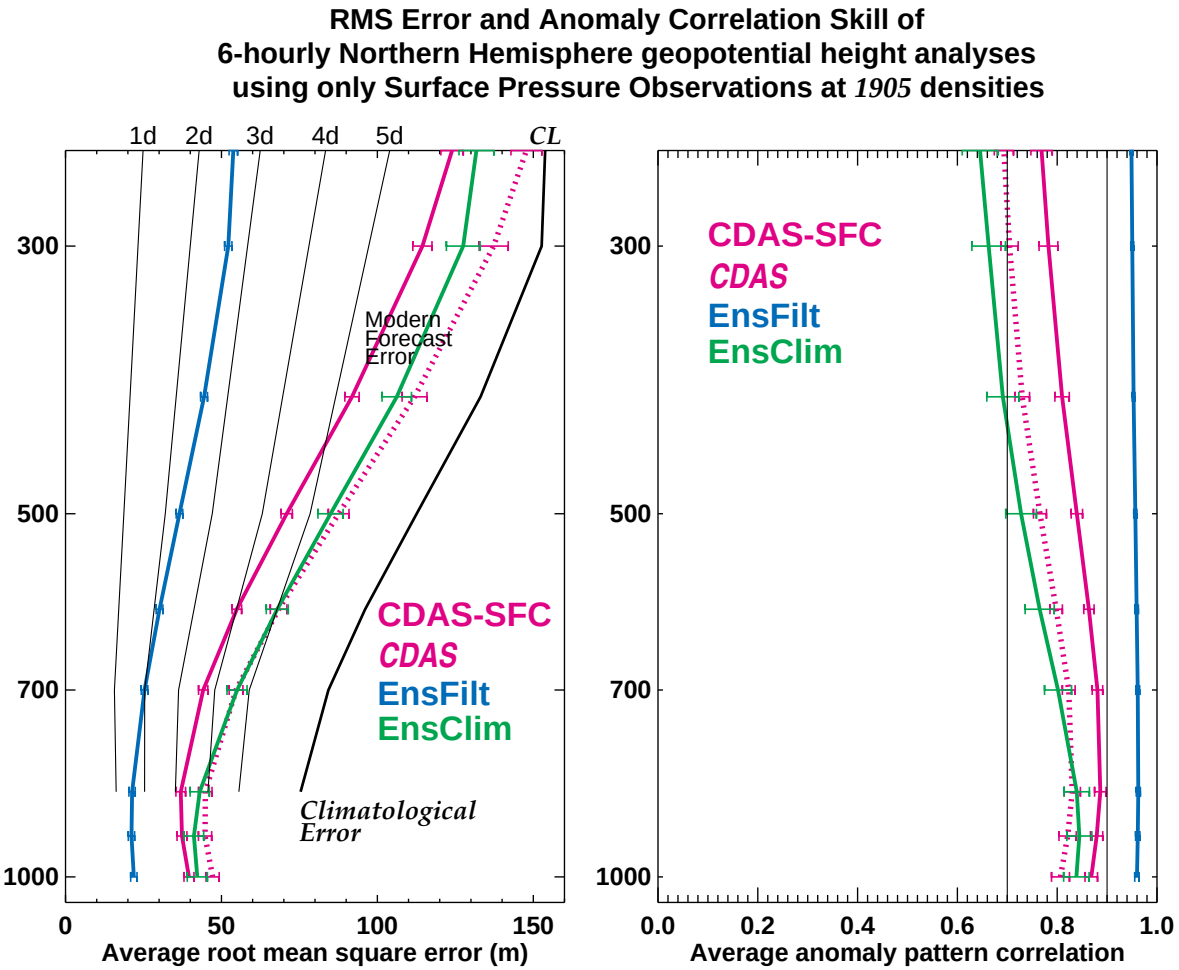

Figure 4: As in Fig. 3 but for simulated 1905 observational densities. Results using the untuned CDAS are indicated by the dotted curves.

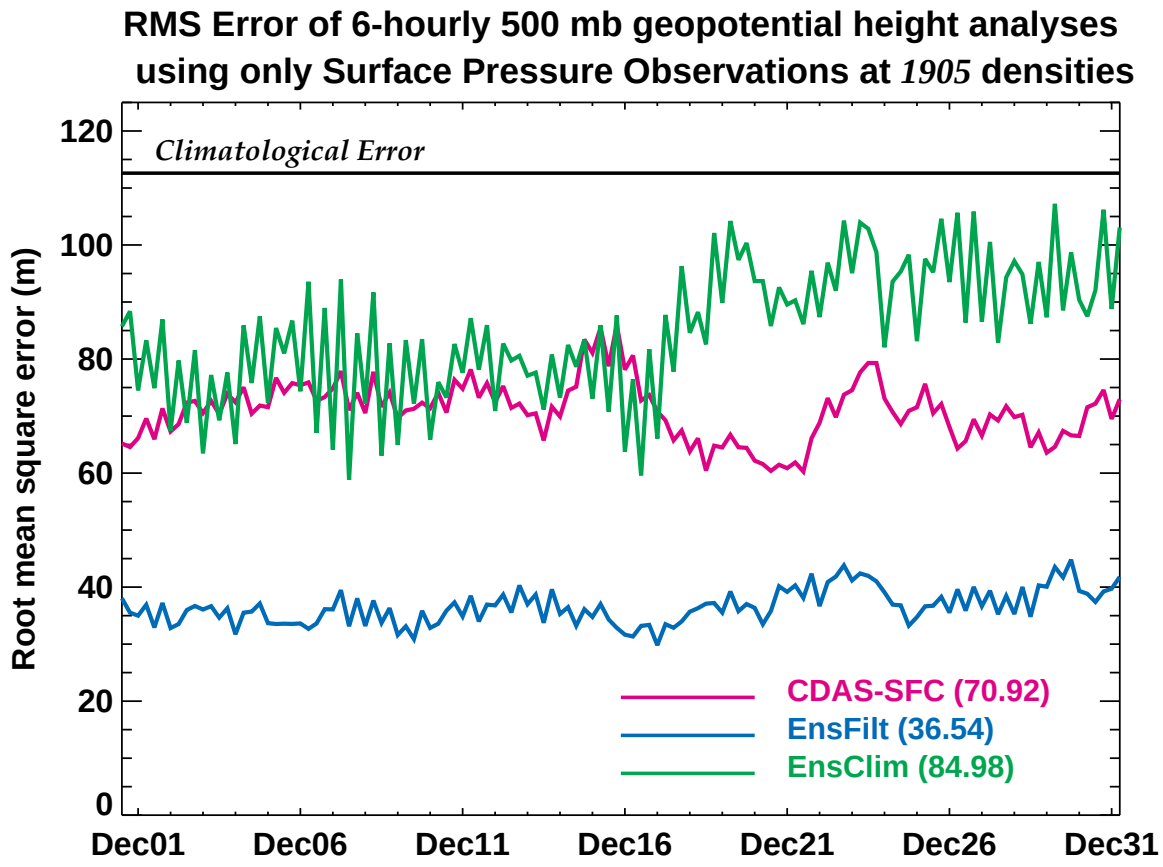

Figure 5. Time series of the 500 hPa geopotential height root mean square difference between analyses of December 2001 from the full NCEP-NCAR reanalysis and from the assimilation experiments. The RMS is averaged over the Northern Hemisphere (20°N-90°N). Climatological standard deviation is indicated by the thick black line.

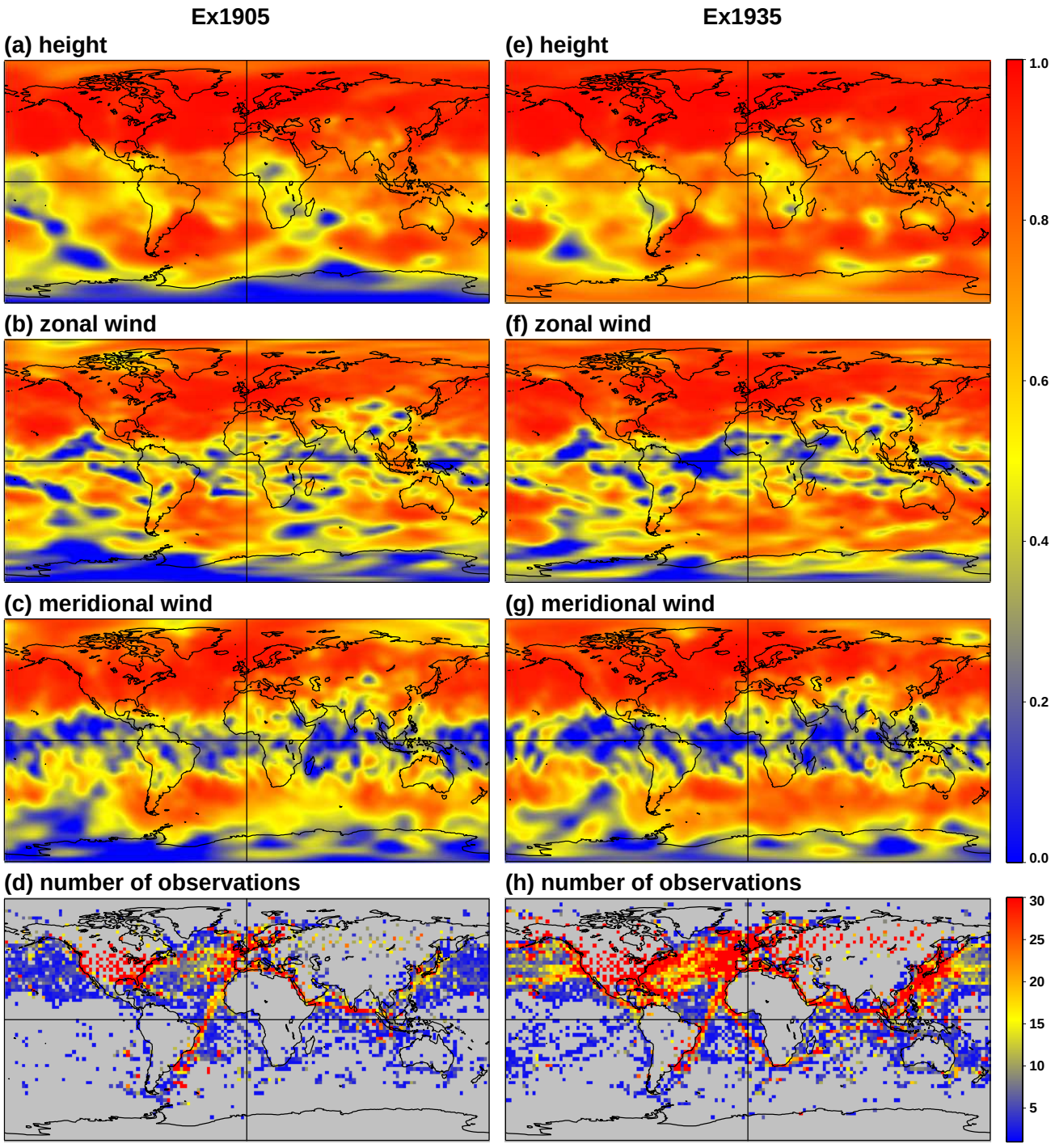

Figure 6. Local anomaly correlation of December 2001 4-times daily analyses from the full NCEP-NCAR reanalysis and (left) 1905 and (right) 1935 assimilation experiments using the Ensemble Filter. Correlations are shown for (a),(e) 700 mb geopotential height; (b),(f) 700 mb zonal wind; and (c),(g) 700 mb meridional wind. Colors in the bottom panels indicate the number of surface pressure observations used in each 2.5° X 2.5° gridbox.

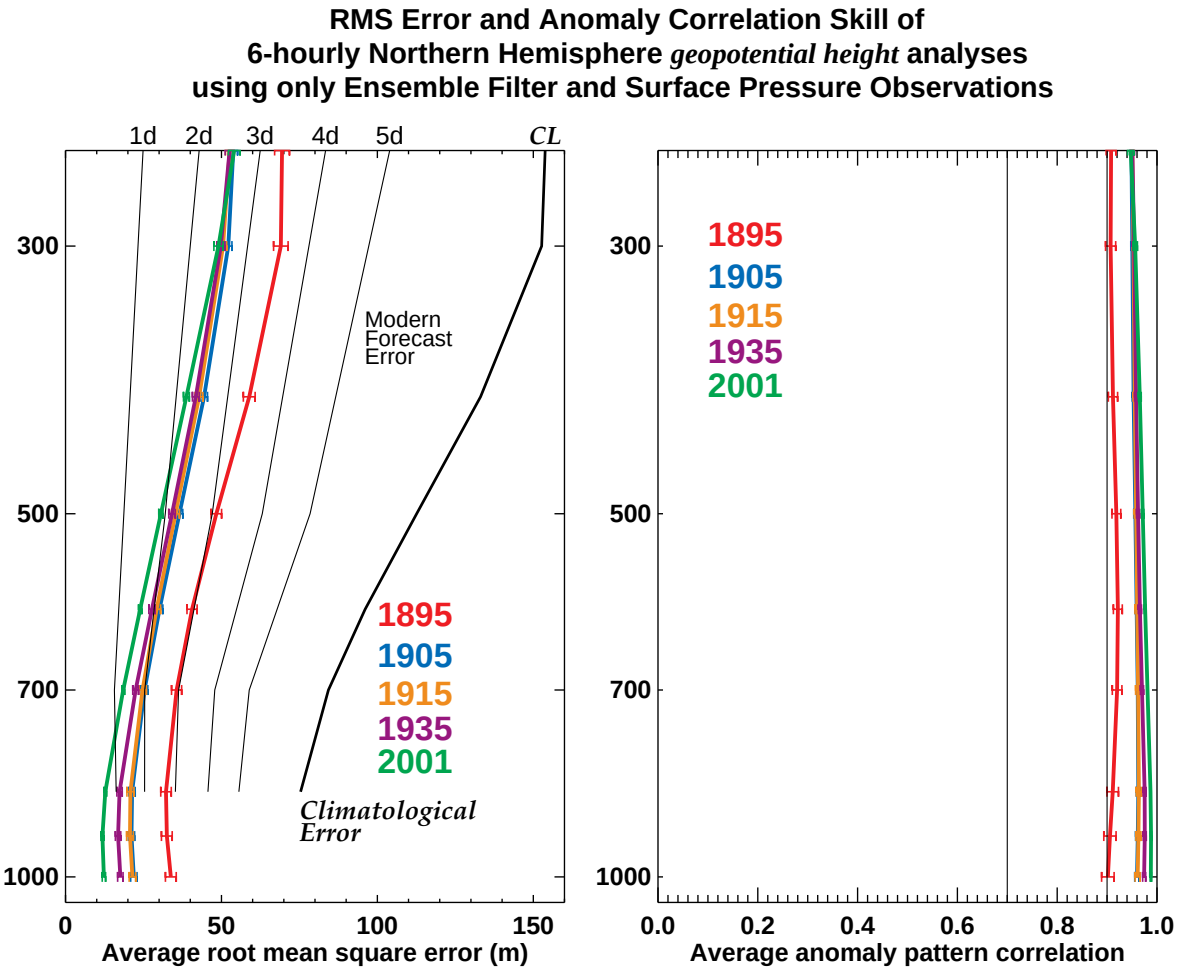

Figure 7. Variation of skill analyzing December 2001 using different simulated networks and the Ensemble Filter. (left) Colored curves show the average root-mean-square (RMS) difference over the Northern Hemisphere (20°N-90°N) between 6-hourly geopotential height analyses from full NCEP-NCAR reanalysis and Ensemble Filter assimilation experiments using only surface pressure observations at 1895, 1905, 1915, 1935, and 2001 densities. Thin solid black curves show *modern* RMS error of NWP forecasts at days 1 through 5. Thick solid black curve shows climatological standard deviation. (right) Colored curves show the monthly average of 6-hourly anomaly pattern correlations over the Northern Hemisphere of geopotential height anomalies from full NCEP-NCAR reanalyses and assimilation experiments.

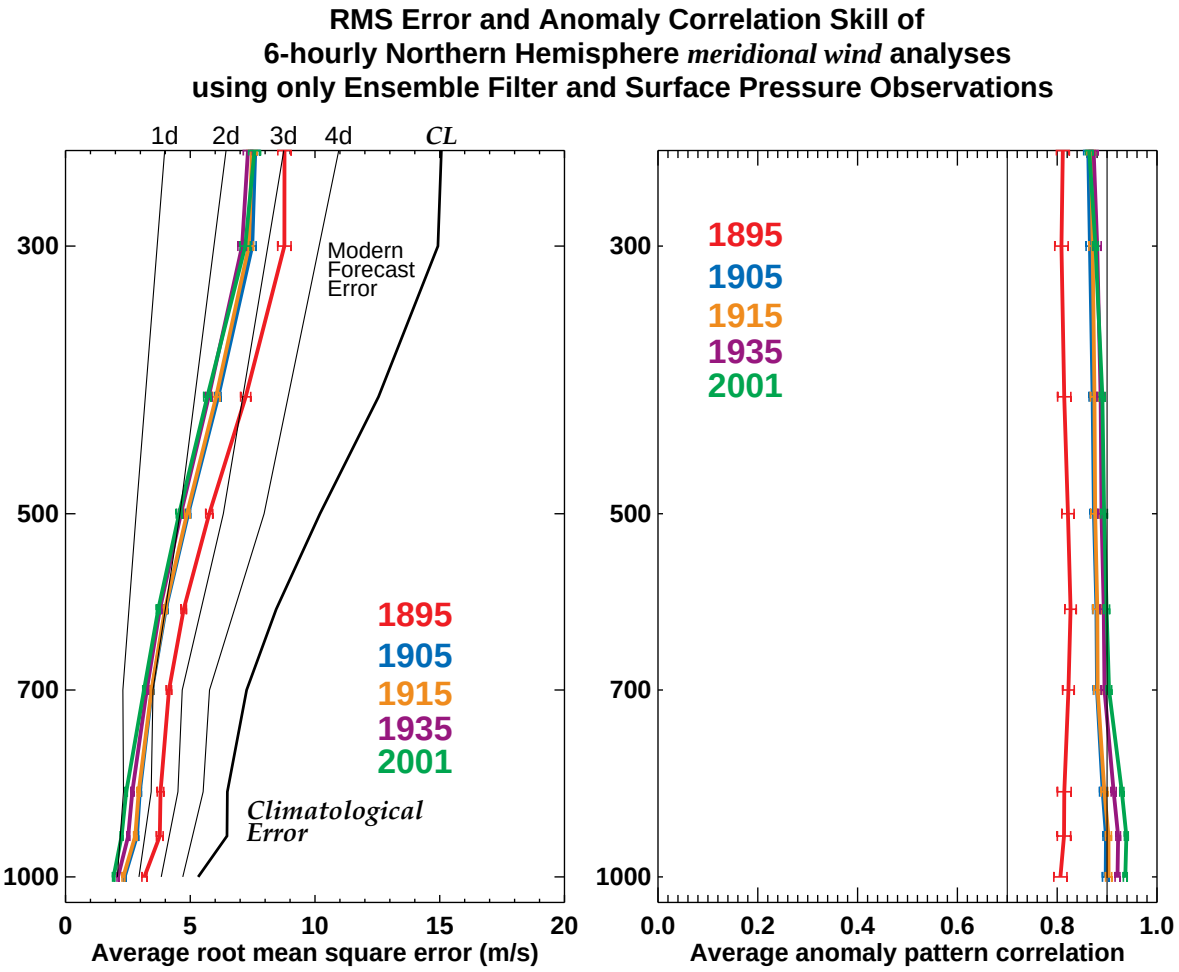

Figure 8. Variation of skill over the Northern Hemisphere analyzing December 2001 meridional wind using varying simulated networks and the Ensemble Filter. As in Fig. 7.

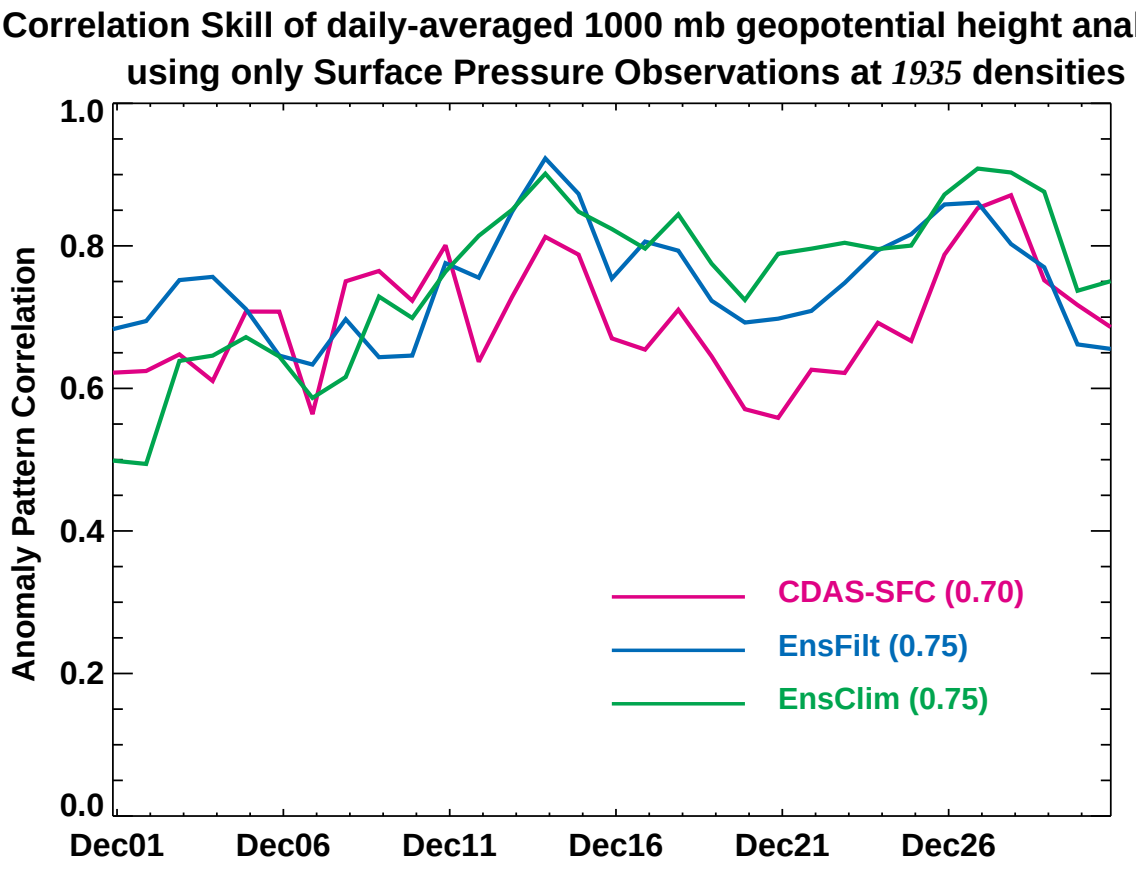

Figure 9. Timeseries of daily-averaged 1000 hPa height anomaly pattern correlation over the Tropics (20°N-20°S). Assimilation experiment analyses of December 2001 using the simulated 1935 surface pressure network are correlated with the full NCEP-NCAR reanalyses.

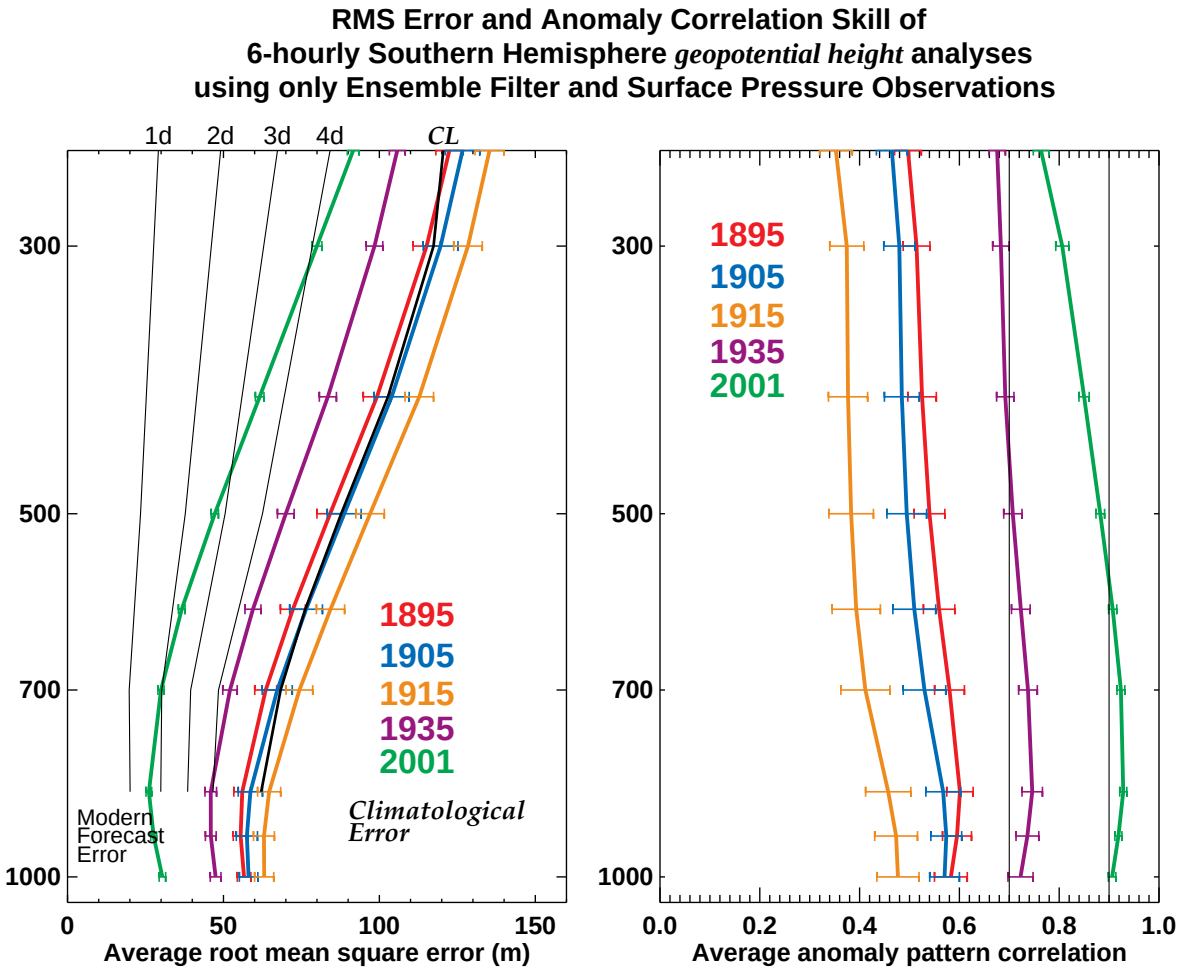

Figure 10. As in Fig. 7 for December 2001 but over the Southern Hemisphere (20°S-90°S).

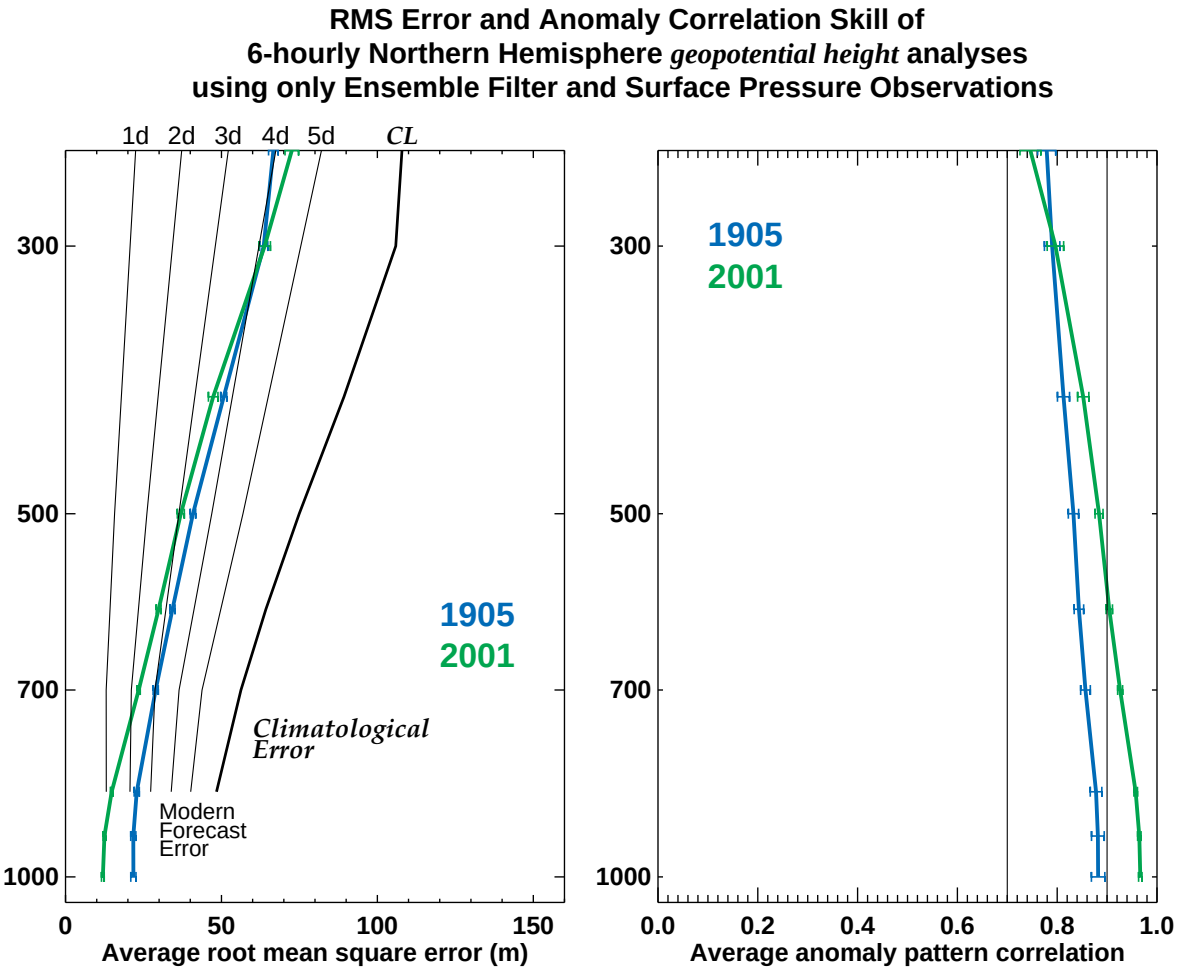

Figure 11. As in Fig. 7 but skill for June 2001 over the Northern Hemisphere using only simulated 1905 and 2001 networks and Ensemble Filter.

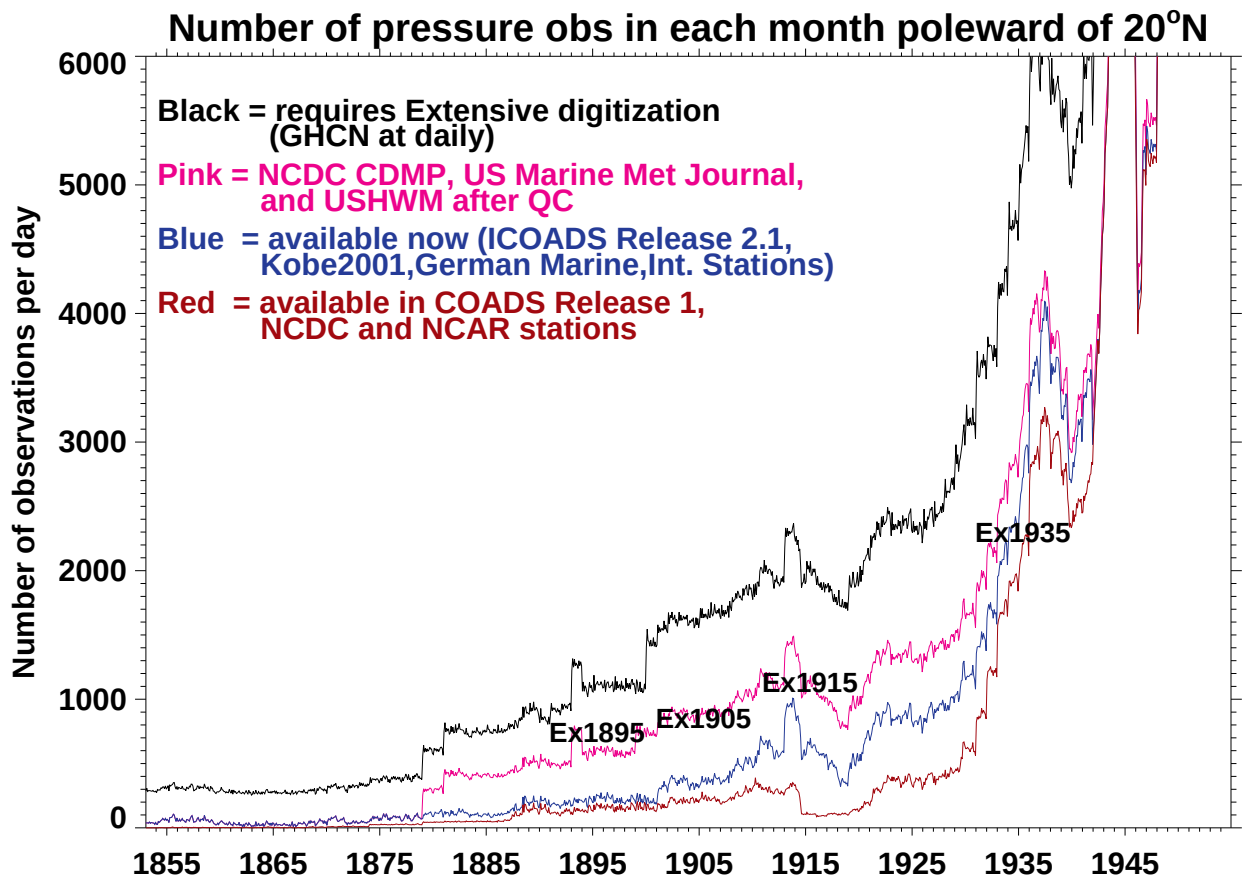

Figure Sidebar 1. Our estimate of the number of Northern Hemisphere (20°N-90°N) surface pressure observations available per day from land and marine sources. Red curve: based on previous COADS and land sources. Blue curve: now available including ICOADS Release 2.1 and Kobe Collection 2000. Pink curve: includes digitization underway, US Marine Meteorological Journals, and quality control of the US Historical Weather Map Series. Black curve: includes international and US data available in manuscript form in NCDC archives. Years indicated with "Ex" show the number of observations used per day in the assimilation experiments.

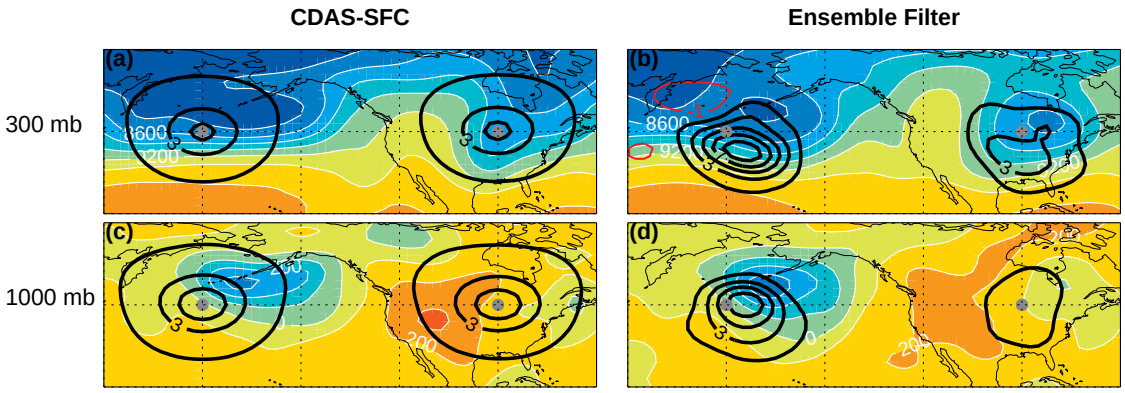

Figure Siderbar 2. Effect of a single surface pressure observation (gray dot) on the final analysis. Filled contours show the first guess field from (a,c) CDAS-SFC and (b,d) Ensemble Filter for geopotential height at (a,b) 300 hPa and (c,d) 1000 hPa. Line contours show the analysis increment after assimilating a single pressure observation that is 1 hPa larger than the first guess pressure field at the indicated location (gray dots). The two gray dots represent two separate experiments. First guess (filled) contour interval is 100 hPa. Analysis increment (line) contour interval is 2 m starting at 1m. Positive (negative) increments are indicated with black (red) contours.